

# Movie Recommendation Systems

---

Nipun Batra

July 21, 2025

IIT Gandhinagar

User/Movie	$P_1$	$P_2$	... ..	$P_M$
$U_1$	3	4	.... ? ....	?
$U_2$	2	?	.....	4
$U_3$	4	5	.....	?
:	:			
:	:			
:	:			
$U_N$	5	1	...?...?...	3

User/Movie	$P_1$	$P_2$	...	$P_M$
$U_1$	3	4	.... ? ....	?
$U_2$	2	?	.....	4
$U_3$	4	5	.....	?
:	:			
:	:			
:	:			
$U_N$	5	1	...?...?...	3

↑ N

← M →

PREDICT

User/Movie	$P_1$	$P_2$	$\dots$	$P_M$
$U_1$	3	4	$\dots$ ? $\dots$	?
$U_2$	2	?	$\dots$ ? $\dots$	4
$U_3$	4	5	$\dots$ ? $\dots$	?
:	:			
$U_N$	5	1	... ? ... ? ..	3

Let's consider a subset of users and movies  
and assume complete data

	Sholay	Swades	Batman	Interstellar	Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮

What can you say about U1?

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮

What can you say about U1? Likes Bollywood  
Dislikes Hollywood

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮

→ What can you say about U2?

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮

→ What can you say about U2? Likes "engineering"

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮

→ What can you say about U3?

Let's consider a subset of users and movies

	Sholay	Swades	Batman	Interstellar	Shawshank
U1	5	5	3	3	2
U2	2	4	5	5	3
U3	2	2	3	3	5
⋮	⋮	⋮	⋮	⋮	⋮

→ What can you say about U3? Likes "shorter" movies

	Sholay	Swades	Batman	Interstellar	The Shawshank
Bollywoodness	1.2	0.95	0.10	0.01	0.02
Engineering	0.1	0.8	0.9	0.95	0.01
Length	0.1	0.08	0.2	0.25	0.90

	Sholay	Swades	Batman	Interstellar	The Shawshank
Bollywoodness	1.2	0.95	0.10	0.01	0.02
Engineering	0.1	0.8	0.9	0.95	0.01
Length	0.1	0.08	0.2	0.25	0.90

\* Describe each movie with some "x" features

\* we have created a matrix  $H$  of size  $x \times M$   
 $C$   
# movies

	Bollywoodness	Engineering	Length
U1	4.0	0.7	0.7
U2	...	...	...
U3	...	...	...

x Describe each user with some "x" features

	Bollywoodness	Engineering	Length
U <sub>1</sub>	4.0	0.7	0.7
U <sub>2</sub>	...	...	...
U <sub>3</sub>	...	...	...
:			

- \* Describe each user with some "x" features
- \* Create a matrix W of size Nusers X x features

Sholay

5

Swades

5

Batman

3

Interstellar

3

Shawshank

2

U1 has rated Sholay 5

Sholay

Swades

Batman

Interstellar

Shawshank

U1

5

5

3

3

2

U1 has rated Sholay

5

Bollywoodness

Engineering

length

Bollywoodness

Sholay

Swades

Batman

U1

4.0

0.7

0.7

Engineering

1.2

0.95

0.01

:

:

:

:

length

0.1

0.08

0.2

:

:

:

:

Sholay

Swades

Batman

Interstellar Shawshank

U1

5

5

3

3

2

U1 has rated Sholay 5

Bollywoodness

Engineering

length

Bollywoodness

Sholay

Swades

Batman

U1

4.0

0.7

0.7

Engineering

1.2

0.95

0.01

:

:

:

:

length

0.1

0.08

0.2

Sholay

Swades

Batman

Interstellar

Shawshank

U1

5

5

3

3

2

U1 has rated Sholay 5

Bollywoodness

Engineering

length

Bollywoodness

Sholay

Swades

Batman

U1

4.0

0.7

0.7

Engineering

1.2

0.95

0.01

:

:

:

:

length

0.1

0.08

0.2

Sholay

Swades

Batman

Interstellar

Shawshank

U1

5

5

3

3

2

U1 has rated Sholay 5

Bollywoodness

Engineering

length

U1

4.0

0.7

0.7

Sholay

1.2

Swades

0.95

Batman

0.01

:

:

:

:

Bollywoodness

Engineering

length

0.1

0.8

0.08

0.9

0.2

Sholay

Swades

Batman

Interstellar

Shawshank

U1

5

5

3

3

2

U1 has rated Sholay 5

Bollywoodness

Engineering

length

U1

4.0

0.7

0.7

Sholay

1.2

Swades

0.95

Batman

0.01

:

:

:

Engineering

0.1

Swades

0.8

Batman

0.9

:

:

:

length

0.1

Sholay

0.08

Batman

0.2

$$4 \times 1.2 + 7 \times 1 + 7 \times 1 \approx 5$$

# MATRIX FACTORISATION

$$A_{N \times M} \approx W_{N \times r} H_{r \times M}$$

- \*  $r \ll N$  and  $r \ll M$
- \* Also called low rank decomposition

# MATRIX FACTORISATION

$$A_{N \times M} \approx W_{N \times r} H_{r \times M}$$

\* Goal: given  $A$ , learn  $W$  and  $H$  s.t.

$$A \approx WH$$

\* or,

$$W^*, H^* = \underset{W, H}{\operatorname{argmin}} \| (A - WH) \|_F^2$$

A side : NORMS

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$\|y\|_2^2 = ?$$

A side : NORMS

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$\|y\|_2^2 = ?$$

Above is square of  $\ell_2$  norm of  $y$

$$= 1^2 + 2^2 + 3^2 + 4^2 = 30$$

A side : NORMS

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\|A\|_F^2 = ?$$

A side : NORMS

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\|A\|_F^2 = ?$$

↑  
Frobenius norm

$$\begin{aligned} &= \sum_{i=1}^3 \sum_{j=1}^4 |a_{ij}|^2 = 1^2 + 1^2 + 2^2 + 2^2 \\ &\quad + 2^2 + \dots \\ &\quad + \dots + 1^2 \end{aligned}$$

## MATRIX FACTORISATION

$$W^*, H^* = \underset{W, H}{\operatorname{argmin}} \| (A - WH) \|_F^2$$

Q: How to learn W and H

# MATRIX FACTORISATION

$$W^*, H^* = \underset{W, H}{\operatorname{argmin}} \| (A - WH) \|_F^2$$

## METHOD I (Gradient Descent)

1) INIT  $W$  and  $H$  as  $N \times r$  and  $r \times N$  matrices

2) FOR  $i$  in  $[1, \dots, \text{ITER}]$ :

$$W = W - \alpha \frac{\partial \| A - WH \|_F^2}{\partial W}$$

$$H = H - \alpha \frac{\partial \| A - WH \|_F^2}{\partial H}$$

# MATRIX FACTORISATION

$$W^*, H^* = \underset{W, H}{\operatorname{argmin}} \| (A - WH) \|_F^2$$

## METHOD I (Gradient Descent)

1) INIT W and H as  $N \times r$  and  $r \times N$  matrices

2) FOR  $i$  in  $[1, \dots, \text{ITER}]$ :

$$W = W - \alpha \frac{\partial \| A - WH \|_F^2}{\partial W}$$

$$H = H - \alpha \frac{\partial \| A - WH \|_F^2}{\partial H}$$

## METHOD II (Alternating least squares)

1) INIT W

2) Till convergence

- Fix W and learn

- H via least sq.

- Fix H and learn

- W via least sq.

Alternating least squares (Intro, Rest in assignment)

$$A = \begin{bmatrix} & \\ & \\ & \\ & \end{bmatrix}_{N \times M}$$

$$w = \begin{bmatrix} & \\ & \\ & \\ & \end{bmatrix}_{N \times r}^n \begin{bmatrix} & \\ & \\ & \\ & \end{bmatrix}_{n \times M}$$

Alternating least squares (Intro, Rest in assignment)

$$A = \begin{bmatrix} & \\ & \end{bmatrix}_{N \times M} \approx w = \begin{bmatrix} & \\ & \end{bmatrix}_{N \times r} \begin{bmatrix} & \\ & \end{bmatrix}_{r \times M}^n$$

Remember linear regression

$$y = \begin{bmatrix} & \end{bmatrix}_{N \times 1} \approx x \begin{bmatrix} & \\ & \end{bmatrix}_{N \times d} \begin{bmatrix} & \end{bmatrix}_{d \times 1}^\Theta$$

$\hat{\Theta} = LS(x, y)$

Alternating least squares (Intro, Rest in assignment)

$$A = \begin{bmatrix} \text{pink shaded} \\ \vdots \end{bmatrix}_{N \times M} \approx w = \begin{bmatrix} \text{light blue shaded} \\ \vdots \end{bmatrix}_{N \times r} \begin{bmatrix} \text{green shaded} \\ \vdots \end{bmatrix}_r^{n \times M}$$

Remember linear regression

$$y = \begin{bmatrix} \text{pink shaded} \\ \vdots \end{bmatrix}_{N \times 1} \approx x = \begin{bmatrix} \text{light blue shaded} \\ \vdots \end{bmatrix}_{N \times d} \theta = \begin{bmatrix} \text{green shaded} \\ \vdots \end{bmatrix}_d^{1 \times 1}$$

$\hat{\Theta} = LS(x, y)$

Alternating least squares (Intro, Rest in assignment)

$$A = \begin{bmatrix} \text{pink} \end{bmatrix}_{N \times M} \approx w = \begin{bmatrix} \text{light blue} \end{bmatrix}_{N \times r} \begin{bmatrix} \text{green} \end{bmatrix}^T_{r \times M}$$

$$\hat{H}[:, 0] = LS(w, A[:, 0])$$

Remember linear regression

$$y = \begin{bmatrix} \text{pink} \end{bmatrix}_{N \times 1} \approx x = \begin{bmatrix} \text{light blue} \end{bmatrix}_{N \times d} \theta = \begin{bmatrix} \text{green} \end{bmatrix}_{d \times 1}$$

$$\hat{\theta} = LS(x, y)$$